Estimate of differential stress in the upper crust from variations in topography and strike along the San Andreas fault

Yuri Fialko,1 Luis Rivera2 and Hiroo Kanamori3

1Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093, USA.
E-mail: fialko@radar.ucsd.edu
2Ecole et Observatoire de Sciences de la Terre, Universite Louis Pasteur-CNRS, Strasbourg, France. E-mail: luis@sismo.u-strasbg.fr
3Seismological Laboratory, California Institute of Technology, Pasadena, CA 91125, USA. E-mail: hiroo@caltech.edu

Accepted 2004 October 15; Received 2004 September 15; in original form 2004 April 21

SUMMARY

The major bends of the San Andreas fault in California are associated with significant variations in the along-fault topography. The topography-induced perturbations in the intermediate principal stress may result in the rotation of the fault with respect to the maximum compression axis provided that the fault is non-vertical, and the slip is horizontal. The progressive fault rotation may produce additional topography via thrust faulting in the adjacent crust, resulting in a positive feedback. The observed rotation of the fault plane due to the along-fault variations in topography is used to infer the magnitude of the in situ differential stress. Our results suggest that the average differential stress in the upper crust around the San Andreas fault is of the order of 50 MPa, implying that the effective fault strength is about a factor of two lower than predictions based on Byerlee’s law and the assumption of hydrostatic pore pressure.

Key words: crustal deformation, fault tectonics, rock fracture, San Andreas fault, seismotectonics, transform faults.

1 INTRODUCTION

The San Andreas fault (SAF) in California is a prime example of a mature transform plate boundary in a continental setting. The 500-km long SAF exhibits several spectacular deviations from planarity (such as the Big Bend north of Los Angeles, Fig. 1), which are associated with variations in local topography. The observed correlation between the fault bends and topography is usually explained in terms of the kinematics of a long-term fault motion: the restraining fault bends result in a local compression, and ‘push-up’ thrust faulting, while the releasing bends produce extension, and ‘pull-apart’ normal faulting (Twiss & Moore 1992; Davis & Reynolds 1996). However, the relationships between the fault geometry, topography, and state of stress in the brittle crust have not been explicitly considered. In this paper we investigate the dynamics of a long-term slip on non-optimally oriented faults in the presence of variations in the along-fault topography. We use geological and seismological constraints on the present-day configuration of the SAF in the areas of major fault bends to derive a relationship between an increase in the vertical stress due to an excess topography, and the fault curvature. Our model is then used to estimate the magnitude of average differential stress in the brittle crust around the SAF.

2 DATA AND THEORY

The topography around the Big Bend and the San Bernardino segments of the SAF is shown in Figs 2(a) and (c), respectively. As can be seen from Fig. 2, there is a significant correlation between the fault strike and local topography. The fault rotates by about 20°–30° where it passes through an elevated terrain. We propose that this rotation is consistent with the topography-induced increase in vertical stress at the seismogenic depth, and may be used to infer the magnitude of the regional differential stress.

According to the Mohr–Coulomb theory, shear failure in a homogeneous medium subject to a uniform triaxial stress should occur on optimally oriented planes (i.e. planes containing the intermediate principal stress axis, and oriented at an acute angle to the maximum compressive stress axis) (Anderson 1951; Scholz 1990; King et al. 1994). Failure occurs when the differential stress $\sigma_1 - \sigma_3$, where $\sigma_1$ and $\sigma_3$ are the maximum and minimum compressive stresses, respectively, reaches a critical value of the order of the effective normal stress (normal stress minus the pore pressure) (Byerlee 1978). The Mohr–Coulomb failure criterion is intrinsically 2-D, and independent of the value of the intermediate stress $\sigma_2$. Assuming that the intermediate stress axis is oriented vertically, as appropriate for strike-slip faulting (Anderson 1951), the Mohr–Coulomb theory...
Figure 1. Map showing the geometry of the San Andreas fault (SAF, thick wavy line) in California. Rectangular boxes outline the major fault bends (SE—San Emigdio bend (‘Big Bend’), SB—San Bernardino bend).

Figure 2. Digital elevation map (Farr & Kobrick 2000) of areas adjacent to the SAF near (a) the Big Bend (San Emigdio bend), and (c) the San Bernardino bend. Figs (b) and (d) show the along-fault variations in topography (triangles) and fault strike (solid lines). Topography data represent elevation above sea level taken from 10-km wide swaths centred on the fault trace (rectangular boxes in parts a and c). The local fault strike is measured clockwise from the north using the same fault discretization as shown in parts a and c.

predicts horizontal slip on vertical planes, regardless of variations in the intermediate stress. There is evidence, however, that the SAF may not be optimally oriented for failure along much of its length (Zoback et al. 1987; Mount & Suppe 1987; Nur et al. 1993), and, furthermore, may have a non-vertical dip angle within the major fault bends while keeping an essentially horizontal slip direction (Nicholson et al. 1986; Griscom & Jachens 1990; Seeber & Armbruster 1995). If so, the mechanics of slip on non-optimally oriented surfaces may be explicitly dependent on the intermediate stress \( \sigma_2 \). Slip on a non-optimally oriented surface may occur, for example if the work required to generate a new fault is greater than the work against friction on a pre-existing surface.

Consider an arbitrarily oriented fault plane defined by a unit normal \( n_i \) \((i = 1, 2, 3)\) in a coordinate system \( x_i \) associated with the principal stresses \((\sigma_1, \sigma_2, \sigma_3)\) (Fig. 3). Hereafter, the intermediate stress \( \sigma_2 \) is assumed to be vertical. Components of the unit vector \( n_i \) satisfy the following identity:

\[
 n_i n_i = n_1^2 + n_2^2 + n_3^2 = 1\
\]

(as usual, the repeating indexes imply summation). Let \( \tau_{ij} \) be the stress tensor, such that in the principal stress coordinate system
The normal stress $\tau_n$ and the shear stress $\tau_s$ resolved on the fault surface are

$$
\tau_n = \tau_i n_i n_i = \sigma_1 n_1^2 + \sigma_2 n_2^2 + \sigma_3 n_3^2,
$$

and

$$
\tau_s = \tau_i n_i j_i.
$$

In eq. (4), $i$ is the unit vector denoting a direction in which the shear stress $\tau_s$ is resolved on a fault plane. For a given stress state, the maximum shear traction $\tau_{\text{m1}}$ acting on a fault plane depends on the fault orientation only:

$$
\tau_{\text{m1}} = \left| \begin{array}{c}
\sigma_1 - \tau_n n_1 \\
\sigma_2 - \tau_n n_2 \\
\sigma_3 - \tau_n n_3
\end{array} \right|
$$

where $\tau_{\text{m1}}$ is the direction in which $\tau_{\text{m1}}$ is acting. Analogous expressions for the normal and shear stresses resolved on a fault plane can be found by considering a 3-D Mohr circle diagram (Jaeger 1956). Eqs (3) and (5) indicate that for vertical faults ($n_2 = 0$) both the normal and the maximum shear stresses (and thus the Mohr–Coulomb failure criterion) are independent of the intermediate stress $\sigma_2$, as expected. We are interested in the dynamics of slip on a fault plane that may not be optimally oriented for failure, and is not collinear with any of the principal stress axes. It is reasonable to assume that the long-term slip on mature faults occurs in the direction of the maximum shear traction $\tau_{\text{m1}}$. This assumption is similar to those used in, for example, inversions of the fault slip or the earthquake focal mechanism data for the principal stress orientations (Gephart & Forsyth 1984; Rivera & Cisternas 1990; Hardebeck & Hauksson 1999). Eq. (5) indicates that for non-vertical faults ($n_2 \neq 0$) the condition of pure strike-slip ($\tau_{\text{m1}}^n = 0$) can be satisfied if and only if the vertical stress equals the normal stress resolved on the fault plane:

$$
\tau_n = \sigma_2.
$$

Condition (6) can be exploited to establish a relationship between variations in the vertical stress $\sigma_2$ (e.g. due to topography), and the associated variations in the fault orientation $n_i$ that render the equilibrium values of the fault-normal stress $\tau_n$. Eliminating an explicit dependence of the normal stress on $n_i$ using constraint (1),

$$
\tau_i n_i n_i = \sigma_1 n_1^2 + \sigma_2 n_2^2 + \sigma_3 n_3^2,
$$

from eqs (3) and (6) one obtains

$$
\sigma_2 - \sigma_3 - (\sigma_1 - \sigma_3) n_1^2 - (\sigma_2 - \sigma_3) n_2^2 = 0.
$$

The relationship (7) admits separation of terms that depend only on the state of stress ($\sigma_1$) and the fault orientation ($n_i$):

$$
\frac{\sigma_2 - \sigma_3}{\sigma_1 - \sigma_3} = \frac{n_2^2}{1 - n_2^2}.
$$

The left-hand side of eq. (8) is a well-known stress ratio $R$ that characterizes the shape of the stress ellipsoid (Bott 1959; Rivera & Cisternas 1990). The right-hand side of eq. (8) can be re-written in terms of the direction cosines $n_1$ and $n_2$ only (e.g. using relation 1):

$$
\frac{n_2^2}{1 - n_2^2} = \frac{1}{1 + (n_3/n_1)^2}.
$$

Note that for non-vertical strike-slip faults the relationship (10) must be satisfied regardless of the fault dip angle.

Provided that the horizontal stress is homogeneous, eq. (10) predicts that an increase in the intermediate stress $\sigma_2$ (e.g. due to an excess topography) should cause an increase in the angle $\phi$, i.e. a rotation of a fault in the horizontal plane away from the $\sigma_1$ axis. Conversely, a decrease in $\sigma_2$ should cause a fault rotation towards the $\sigma_1$ axis. These predictions are consistent with the observed variations in topography and fault strike along the SAF (Fig. 1), and the approximately NNE orientation of the maximum compression axis in California (Lisowski et al. 1991; Unruh et al. 1996). In particular, the SAF rotates counter-clockwise where it crosses the mountains (Big Bend and San Bernardino areas), and clockwise where it drops below the sea level (e.g. north of Point Arena and around the Salton Sea, see Fig. 1). Predictions of eq. (10) admit a simple physical interpretation. An increase in the vertical stress on a dipping fault introduces a component of normal faulting into the resolved shear stress on the fault. In order to maintain the strike-slip, the fault must rotate to higher angles to the maximum compressive stress axis, thereby introducing a component of thrust faulting that will balance the component of normal faulting introduced by the excess topography.

The topography-induced perturbations in the intermediate stress in the upper crust may be readily estimated given the ‘excess’ topography $\Delta h$: $\Delta \sigma_2 = \rho^2 g \Delta h$, where $\rho$ is the rock density, and $g$ is the gravitational acceleration. Because the amount of the topography-induced fault rotation necessarily depends on the average differential stress $\sigma_1 - \sigma_3$ in the brittle layer, the latter may be determined from...
eq. (10) by relating the estimated changes in the vertical stress $\Delta\sigma_2$ to the observed changes in the fault strike $\Delta\phi$:

$$\sigma_1 - \sigma_3 = \frac{2\Delta\sigma_2}{\cos 2\phi - \cos 2(\phi + \Delta\phi)}. \quad (11)$$

Fig. 4 shows the average differential stress calculated using eq. (11) for a range of possible values of the angle $\phi$ between the fault trace and the maximum compression axis, and the fault rotation $\Delta\phi$ of 20° to 30°, as suggested by the geological data (Figs 1 and 2). In our calculations we assume a rock density of $2.8 \times 10^3$ kg m$^{-3}$, and average excess elevation of 1.5 km along the major bends of the SAF (Fig. 2). For comparison, we also show the maximum average differential stress in the brittle layer predicted by Mohr–Coulomb theory under the assumption of hydrostatic pore pressure:

$$\Delta\sigma_{max} = \frac{1}{H} \int_0^H (\sigma_1 - \sigma_3) \, dz = \frac{(\rho_w - \rho_0)gH\mu}{\sqrt{1 + \mu^2}}, \quad (12)$$

where $\rho_w$ is the water density, $H$ is the thickness of the brittle layer, and $\mu$ is the coefficient of friction (see horizontal lines in Fig. 1) (Byerlee 1978; Townend & Zoback 2000).

3 DISCUSSION

There is a long-standing debate in earthquake physics regarding the magnitude of shear stresses on average supported by major seismogenic faults in the Earth's crust (Scholz 1990; Rice 1992; Zoback et al. 1993). The empirical Mohr–Coulomb failure theory, based on a large number of rock fracture experiments, predicts that faults slip when shear stresses in the host rocks are of the order of the lithostatic stress, provided that the pore fluid pressure does not significantly exceed hydrostatic (Anderson 1951; Byerlee 1978). Assuming a typical depth of the seismogenic zone of 12 km, a rock density of $2.8 \times 10^3$ kg m$^{-3}$, and a coefficient of friction of 0.6, the fault-averaged shear stress on the verge of failure is of the order of 10$^6$ Pa. Measurements in deep boreholes seem to indicate that stresses in the upper crust in general obey Byerlee friction, and, moreover, the state of stress in the crust is near the failure envelope even in areas that are not tectonically or seismically active (Townend & Zoback 2000). On the other hand, typical earthquake stress drops of the order of 10$^6$–10$^8$ Pa (Kanamori & Anderson 1975), the common absence of melting on localized fault slip zones (Chester & Chester 1998), the lack of a significant heat flow anomaly (Lachenbruch & Sass 1980), and the nearly orthogonal orientation of the far-field principal compression axis to the SAF (Mount & Suppe 1987) have been used to argue that large crustal faults may operate at relatively low driving stresses of the order of a few tens of megapascals or less. The latter statement constitutes the ‘weak fault’ hypothesis (Rice 1992). Given the geological and seismological constraints on the contemporaneous kinematics of the SAF, the along-fault variations in strike and topography described in the previous section can be used to provide an independent estimate on the level of the in situ deviatoric stress in the crust around the SAF.

A robust inference of the average differential stress from eq. (11) requires knowledge of the orientation of the maximum principal stress $\sigma_1$ (i.e. the angle $\phi$). Assuming that the regional crustal stress is relatively homogeneous on a spatial scale of the order of the wavelength of the major fault bends (i.e. tens of kilometres), a lower bound on $\phi$ is given by the condition that the rotated section of the fault must have the same sense of shear stress as the rest of the fault, implying $\phi < \pi/2 - \Delta\phi$. The precise values of $\phi$ are the subject of a significant debate. Several lines of evidence, including inversions of seismic focal mechanisms, measurements of the long-term geodetic strain, borehole data, and structural geological mapping around the SAF suggest that the regional $\sigma_1$ axis may be oriented at high angles (60°–80°) to the SAF (Mount & Suppe 1987; Lisowski et al. 1991; Zoback et al. 1993). Other interpretations of similar data favour values of $\phi$ scattered around 45° (McNally et al. 1978; Zoback et al. 1980). Hardebeck & Hauksson (1999) proposed that the $\sigma_1$ axis rotates to intermediate (35°–55°) angles within a few tens of kilometres from the fault, based on inversions of a number of the earthquake focal mechanisms (see shaded area in Fig. 4). Their data also show that the orientation of the $\sigma_1$ axis is approximately constant around the San Bernardino bend, consistent with our assumption of a locally homogeneous stress field. As one can see from Fig. 4, for the intermediate values of angles $\phi$ obtained by Hardebeck & Hauksson (1999) our model gives rise to differential stresses of 40–60 MPa, or fault-averaged shear stresses of the order of 20–30 MPa. The inferred stresses are about two to three times lower than those predicted assuming the laboratory values of the coefficient of friction ($\mu = 0.6–0.8$) and hydrostatic pore pressures (see eq. 11 and Fig. 4), but are on the high end of constraints provided by the heat flow measurements (Brune et al. 1969; Lachenbruch & Sass 1980). These results suggest that the strength of the SAF falls between the predictions of the strong (Scholz 2000) and weak (Rice 1992) fault theories. Recently, Townend & Zoback (2004) analysed an earthquake data set similar to that used by Hardebeck & Hauksson (1999), and concluded that the angle between the maximum compressive stress axis and the SAF in southern California is about 60°–70°. Taken at face value, this inference might be interpreted as indicating the high strength of the SAF (Fig. 4). However, the interpretation of Townend & Zoback (2004) implies that the orientation of the $\sigma_1$ axis correlates with the local fault strike. If so, our assumption of the locally homogeneous horizontal stress field is invalid, and eq. (10) cannot be used to deduce the magnitude of the differential stress.

In addition to the assumption of the homogeneous stress state, our estimate of the in situ differential stress relies on the assumption of horizontal slip on a non-vertical fault. The evidence for a
dipping geometry of the SAF around the San Emigdio, San Bernardino, and Point Arena bends includes seismic, magnetic, gravity, and geological observations (McNally et al. 1978; Nicholson et al. 1986; Griscom & Jachens 1990; Seebier & Armbruster 1995; Magistrale & Sanders 1996; Yule & Sieh 2003), although details of the SAF configuration beneath the major bends are still debated (Carena et al. 2004). While the non-vertical fault orientation may initially develop in response to a transpressional or transen- sional stress regime, thereby allowing the fault to accommodate some fault-normal displacements, the seismic and geome- rological data suggest that the current kinematics of the SAF within the fault bends is predominantly strike-slip, and that most of the conver- gence/extension across the bends is taken up by the thrust/normal faulting in the ambient crust (Twiss & Moore 1992; Davis & Reynolds 1996; Yule & Sieh 2003). The partitioning of oblique convergence onto the purely strike-slip and dip-slip fault systems is well known and has been previously explained in terms of the least-work principle (Michael 1990). However, to date no mechani- cal energy calculations exist for the case of non-planar faults, and it remains to be seen whether the least-work principle can explain the observed geometry of the SAF. We note that the topography-induced fault rotation required to preserve the strike-slip setting decreases shear stress, and increases normal stress on the fault plane, moving it further from the failure envelope. Therefore the long-term per- sistence of the major fault bends implies either that the bends are weaker compared with the rest of the fault, or that the bend geometry minimizes the total work (including the mechanical energy, as well as changes in the gravitational potential energy) spent on slip on the SAF and the adjacent faults in the presence of the along-strike variations in topography. Because the total work done against the fault friction depends on the depth-averaged normal stress within the brittle layer, our model implies that the depth to the brittle–ductile transition does not track the surface topography. The seismic data (in particular, the hypocentral depths of local earthquakes) do not indicate any shallowing of the brittle–ductile transition beneath the major fault bends (Hill et al. 1990; Richards-Dinger & Shearer 2000), lending support to our model assumptions.

We point out that our analysis may be relevant for long-term slip on faults with large total offsets. Individual earthquakes and young faults may exhibit significant deviations from the suggested relationship between the topography and the fault geometry. For example, the surface rupture of the 1999 Hector Mine earthquake in southern California changes its strike where it crosses the relatively small (average elevation of ~0.5 km) Bullion mountains (USGS et al. 2000; Fialko et al. 2001). The sense of the rupture rotation, however, is opposite to that predicted by eq. (10). Time-dependent inversions of seismic data (Ji et al. 2002) indicate that the Hector Mine rupture initiated on a central segment of the fault that may have been more optimally oriented for failure, and then propagated bi-laterally on pre-existing faults that are at a higher angle with respect to the regional compression axis (Nur et al. 1993). Therefore the along-fault variations in strike in the case of the Hector Mine earthquake may be due to the heterogeneous strength of the crust and the dynamic aspects of the rupture, rather than to the long-term effects of topography considered in this paper.

As the along-strike variations in topography affect the fault ge- ometry only insofar as they affect the shear stress resolved on the fault plane, the vertical segments of strike-slip faults should not exhibit any correlation between the local topography and the fault strike. A possible example may be a stretch of the SAF north of the San Bernardino bend (Fig. 2c). This part of the SAF is believed to have a near-vertical dip throughout the brittle layer (Griscom & Jachens 1990), and it does not exhibit any appreciable deviations from planarity where it passes through an elevated terrain (Figs 2c and d). While the particular mechanisms leading to the development of major fault bends remain poorly understood, our model suggests that a relationship between the topographically induced perturbation in the intermediate principal stress, and changes in the fault strike for non-vertical faults may result in a positive feedback mechanism, whereby a build-up in elevation due to thrust faulting around a compressive bend is likely to cause a further rotation of the fault. The long-lived fault bends may therefore represent a stable configuration corresponding to either the reduced effective coefficient of friction within the bent fault section, or the minimum energy dissipation on the entire fault system including the main strike-slip fault, as well as the associated dip-slip faults in the ambient crust. Alternatively, a continued build-up of topography by thrust faulting may result in a complete locking of the fault bend, and a subsequent re-organization of the fault geometry. This may have a bearing on the observed branching of the SAF, for example in the vicinity of the San Bernardino segment (Fig. 2c), and the strain transfer from the SAF system to the Eastern California shear zone (Dokka & Travis 1990; Nur et al. 1993; Rockwell et al. 2000).

4 CONCLUSIONS

We combined the kinematic constraints on the configuration and long-term slip of the San Andreas fault with the dynamics of slip on an arbitrarily oriented surface to investigate the observed correlation between the major bends of the SAF in California and variations in the along-fault topography. We demonstrate that the topography- induced perturbations in the intermediate (vertical) principal stress may enhance bending of a non-optimally oriented fault due to an azimuthal rotation of the pure strike-slip vector. In the presence of the excess topography, the slip vector is predicted to rotate away from the maximum compression axis. The observed rotation of the fault plane due to variations in topography along the SAF is consistent with the model predictions. The fault rotation may result from either the progressive reduction in the effective coefficient of friction within the bent fault section, or the minimization of total work spent on slip on the main strike-slip fault, as well as on the associ- ated dip-slip faults in the ambient crust. Regardless of the particular mechanism of the fault bending, the observed variations in the fault strike and topography may be used to infer the magnitude of the in situ differential stress provided that the following conditions are met: (1) the fault is non-vertical; (2) the long-term fault slip is hori- zontal; and (3) the horizontal stress state has little variability on the spatial scale of the order of the fault-bend wavelength. Several lines of geological and geophysical data indicate that these conditions are indeed satisfied around the major fault bends of the SAF. For the seismologically inferred angles between the average fault strike and the maximum compression axis in southern California of 35°–55° (Hardebeck & Hauksson 1999), our results suggest that the average differential stress in the upper crust is of the order of 50 MPa, im- plying that the effective strength of the San Andreas fault is about a factor of two lower compared with predictions based on Byerlee’s law and the assumption of hydrostatic pore pressures, but is at the high end of constraints provided by heat flow measurements.

ACKNOWLEDGMENTS

We thank two anonymous reviewers and the Editor Cindy Ebinger for their insightful comments and suggestions. This material is based
upon work partially supported by the National Science Foundation under Grant EAR-0338061, and the Southern California Earthquake Center (SCEC). SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 806.

REFERENCES


© 2004 RAS, GJI, 160, 527–532